

## Classificador de máxima verossimilhança aplicado à identificação de espécies nativas na Floresta Amazônica

Luiz Otávio Moras Filho <sup>1</sup>  
Evandro Orfanó Figueiredo <sup>2</sup>  
Marcos Antônio Isaac Júnior <sup>3</sup>  
Vanessa Cabral Costa de Barros <sup>4</sup>  
Marcos Cicarini Hott <sup>5</sup>  
Luís Antônio Coimbra Borges <sup>6</sup>

<sup>1, 3, 4, 5, 6</sup> Universidade Federal de Lavras - UFLA

Caixa Postal 3037 - CEP 37200-000 - Lavras - MG, Brasil

<sup>1</sup> lomf22@gmail.com; <sup>3</sup> maisaacjr@yahoo.com.br; <sup>4</sup> vanessacabralcb@gmail.com;

<sup>5</sup> cicarini@yahoo.com.br; <sup>6</sup> luis.borges@dcf.ufla.br

<sup>2</sup> Empresa Brasileira de Pesquisa Agropecuária Ministério da Agricultura, Pecuária e Abastecimento - Embrapa

Rodovia BR-364, Km 14 - Caixa Postal 321 - CEP: 69900-970 - Rio Branco - AC, Brasil

<sup>2</sup> evandro.figueiredo@embrapa.br

**Abstract.** Among a variety of digital classification methods based on remote sensing images, the Maximum Likelihood (ML) is widely used in environmental studies, mainly for land cover and vegetation analysis. This study aimed to evaluate the effectiveness of supervised classification by ML technique in a forest management area of dense ombrophilous forest, using one RapidEye image. With this purpose, it was conducted the census of species over 30 cm in diameter at breast height and calculated the Cover Value Index (CVI), and selected the 20 species with the highest CVI as a parameter for classification in a Geographic Information System. 13 of the 20 species selected in the study area were not identified by the classification method, and among the seven identified species, two were underestimated and the others were overestimated. Both the maximum likelihood technique and the spatial resolution of the image used were not suitable for supervised classification of native vegetation, with Kappa index of 0.05 and global accuracy of 5.53%. Studies using spectral characterization in leaf level supported by higher or hyper spectral and spatial resolution images are recommended to increase the accuracy of classification.

**Palavras-chave:** remote sensing, image processing, supervised classification, plant cover index, sensoriamento remoto, processamento de imagens, classificação supervisionada, índice de cobertura vegetal.

### 1. Introdução

Os Sistemas de Informações Geográficas (SIG) têm sido amplamente utilizados em estudos ambientais devido a sua capacidade de representar e analisar a territorialidade dos fenômenos nele representados. Esses sistemas consideram o ambiente como uma entidade que tem expressão espacial a ser modelada segundo sua viabilidade taxonômica e julgando como relevante à distribuição territorial das classes de fenômenos nele identificadas (MOURA, 2003; MEIRELLES et al., 2007; MORAS FILHO, 2015).

Para caracterizar uma área de interesse, podem ser aplicadas técnicas de classificação digital, que consistem na implementação de uma regra de decisão para que o computador possa atribuir uma determinada classe a certo conjunto de células (*pixels*) em uma imagem. O objetivo dessas técnicas, segundo Novo (1992) é tornar o processo de mapeamento ou reconhecimento de características da superfície terrestre menos subjetivo e com maior potencial de repetição em situações subsequentes.

De acordo com Jensen (2005), em estudos ambientais, o método mais utilizado para analisar dados e extrair informações refere-se à classificação digital fundamentada em dados multiespectrais, podendo ser realizada de forma supervisionada e não-supervisionada. Nesse trabalho será destacada a classificação supervisionada.

A classificação supervisionada pode ser definida como o processo de usar amostras de identidade conhecida para classificar *pixels* de identidade desconhecida. Neste tipo de classificação, o analista tenta identificar áreas específicas que representem exemplos homogêneos dessas feições (RIBEIRO, 2004). Entre as técnicas de classificação supervisionada mais utilizadas, destaca-se a regra de decisão de Máxima Verossimilhança - *Maximum Likelihood* (ML).

A ML, basicamente, estabelece a probabilidade de uma célula pertencer a cada um dos conjuntos de classes pré-definidos, sendo atribuída à célula a classe de mais elevada probabilidade (NOVO, 1992). É um classificador amplamente utilizado para classificação de cobertura do solo e da vegetação, podendo auxiliar, segundo Beltrame et al. (2007) na identificação de diferentes espécies vegetais dentro do mesmo bioma.

Diante disso, o presente trabalho consiste na avaliação da eficiência do classificador ML para espécies nativas de Floresta Ombrófila Densa em uma área de manejo florestal. Esses dados poderão ser úteis para comparação de estudos relacionados à classificação de espécies nativas, podendo ser utilizados para estimular políticas públicas para o desenvolvimento de uma base georreferenciada de espécies nativas na região.

## 2. Metodologia de Trabalho

Em uma área de manejo florestal localizada entre o par de coordenadas 67°38'81"W, 9°53'39"S e 67°33'23"W, 9°57'33"S, situado no município de Rio Branco, estado do Acre, Brasil, foi realizado censo das espécies florestais com diâmetro na altura do peito (DAP) maior que 30 cm. Os dados foram georreferenciados e exportados para um Sistema de Informações Geográficas (SIG) no formato *shapefile*.

Foram identificados 26.375 indivíduos de 28 famílias e 60 gêneros distintos. Foi calculado o Índice de Valor de Cobertura- *Cover Value Index* (CVI), de acordo com Martins (1991), e selecionada as 20 espécies com maior CVI como testemunha para análise de eficiência da classificação digital.

Realizou-se então a classificação supervisionada baseada na técnica de máxima verossimilhança implementada a um Sistema de Informações Geográficas. Como parâmetro para classificação, foram utilizadas as espécies com maior CVI, denominadas “regiões de interesse”. Foi utilizada uma imagem RapidEye (resolução espacial de 5 m) de 01/09/2014. Os dados foram exportados em formato de texto, tabulados e foram gerados gráficos e tabelas com as informações de cada célula da imagem classificada.

## 3. Resultados e Discussão

Foram selecionados 19.270 indivíduos de 20 diferentes espécies. Com base no valor espectral das células que continham cada um desses indivíduos, a imagem RapidEye foi classificada e geraram-se 647.296 células representando os possíveis indivíduos com maior CVI identificados por sensoriamento remoto na região de estudo. A distribuição das espécies inventariadas e espécies classificadas são representadas na Figura 1.

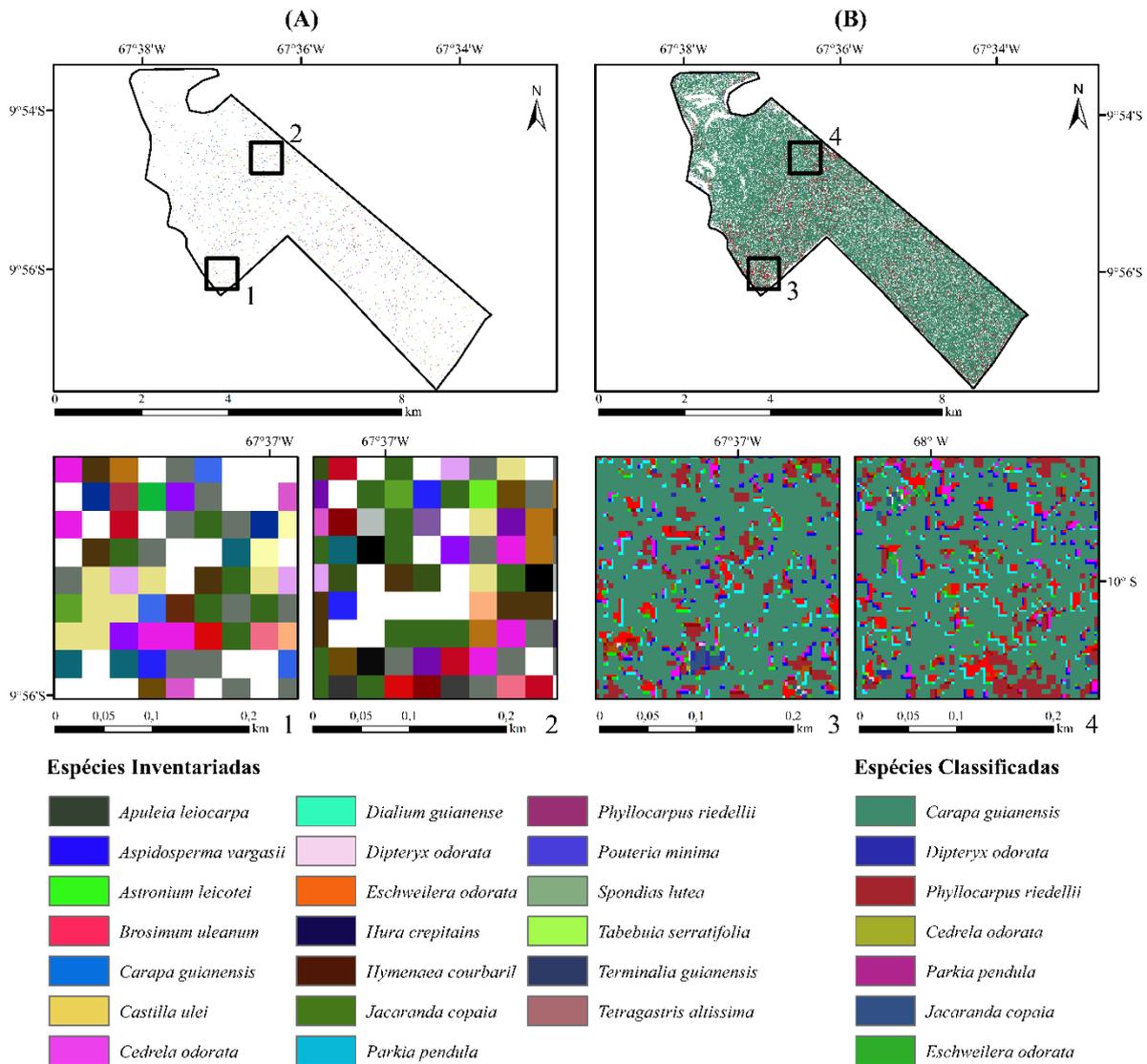


Figura 1. Mapa de localização das espécies inventariadas (A) e de classificação supervisionada por meio do método de Máxima Verossimilhança (B).

A metodologia e os dados utilizados não permitiram a identificação de 13 das 20 espécies selecionadas (*Apuleia leiocarpa*, *Aspidosperma vargasii*, *Astronium leicotei*, *Brosimum uleanum*, *Castilla ulei*, *Dialium guianense*, *Hura crepitans*, *Hymenaea courbaril*, *Pouteria mínima*, *Spondias lutea*, *Tabebuia serratifolia*, *Terminalia guianensis* e *Tetragastris altissima*).

Entre as espécies classificadas, foram subestimadas a classificação das espécies *Cedrela odorata* e *Parkia pendula*, e superestimadas a classificação das espécies *Eschweilera odorata*, *Dipteryx odorata*, *Jacaranda copaia*, *Phyllocarpus riedellii* e *Carapa guianense*. A comparação do número de células que representam os indivíduos selecionados no inventário florestal e da imagem classificada é apresentada na Figura 2.

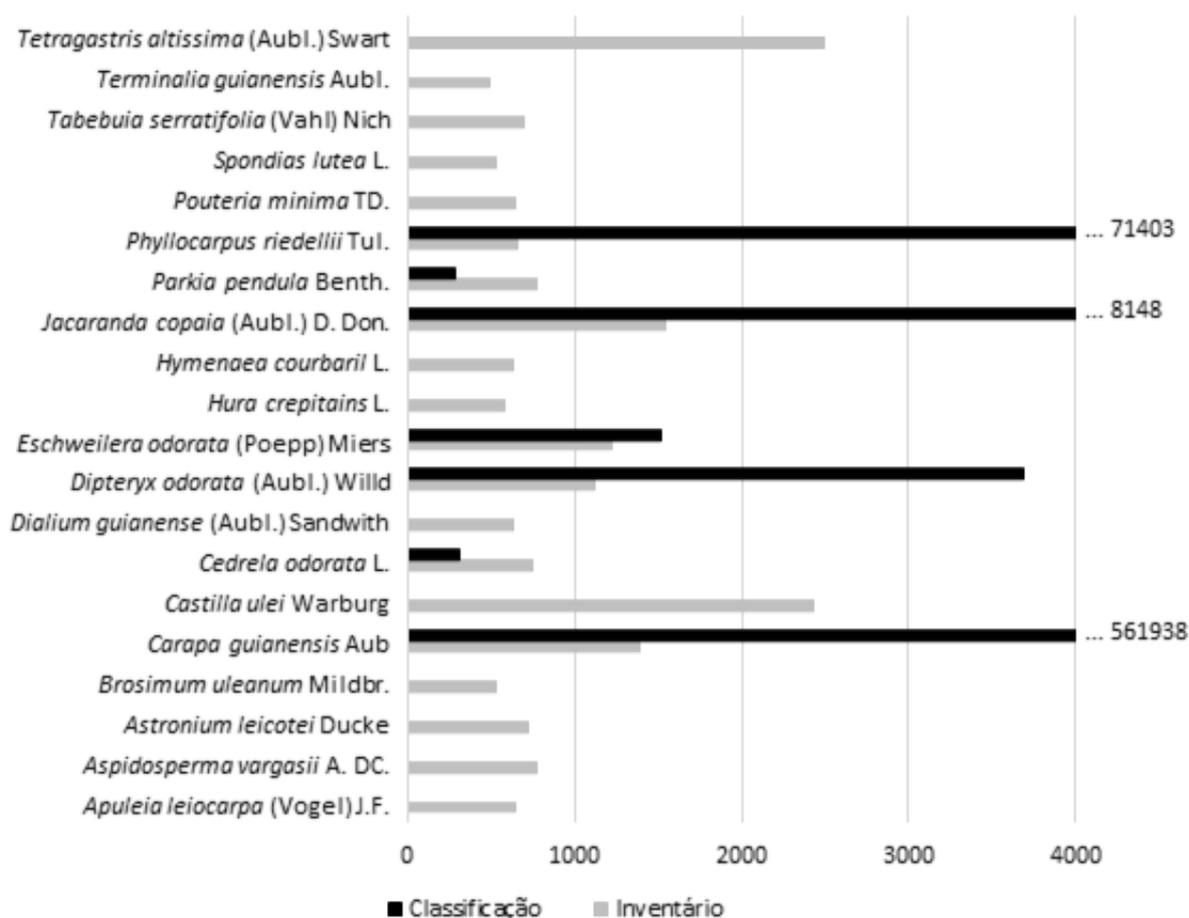


Figura 2. Diferença entre o número de células que representam as espécies inventariadas em campo e as possíveis espécies identificadas no método de classificação supervisionada por técnica de Máxima Verossimilhança.

Apesar das espécies *Tetragastris altíssima* e *Castilla ulei* possuírem os maiores CVI, a metodologia e os dados utilizados foram insuficientes para identifica-las. Na classificação supervisionada por ML dessas espécies florestais, obteve-se um índice Kappa de 0,05 e uma confiança global de 5,53%, calculados por meio de matriz de confusão, em que os erros de omissão e comissão de cada espécie identificada são representados na Tabela 1.

Celinski e Zimback (2010), ao utilizar a técnica de ML para classificação de cobertura vegetal na região do da região dos Campos Gerais, Paraná, tendo como base uma imagem do satélite CBERS-2, detectaram baixa acurácia para classes de eucalipto, sorgo/milho, pasto, soja e feijão. Os autores obtiveram melhores resultados por meio da técnica de Árvore de Decisão.

Avaliando os resultados do mapeamento do uso da terra, vegetação e impactos ambientais realizado por Ortiz e Freitas (2005), no município de Rio Claro, São Paulo, detectou-se uma diferença de, aproximadamente, 26% na classe de silvicultura, 67% para agricultura e 73% para vegetação natural, quando comparadas com a levantamento de verdade terrestre.

Em estudo realizado por Chagas et al. (2003), no mapeamento de classes de solo no nordeste do Rio de Janeiro, foi identifica como uma das mais expressivas fontes de erro a heterogeneidade da área o que, de maneira análoga, é também identificado no presente trabalho.

Tabela 1 – Acurácia da classificação supervisionada por Máxima Verossimilhança.

Espécies	<i>C. guianensis</i>	<i>C. odorata</i>	<i>D. odorata</i>	<i>E. odorata</i>	<i>J. copaia</i>	<i>P. pendula</i>	...	<i>P. riedellii</i>
<i>A. leiocarpa</i>	443	0	4	0	4	1		51
<i>A. vargasii</i>	524	1	3	3	3	0		65
<i>A. leicotei</i>	476	0	4	0	5	1		50
<i>B. uleanum</i>	349	1	4	1	4	0		41
<i>C. guianensis</i>	978	0	9	1	9	1		91
<i>C. ulei</i>	1639	1	16	4	16	1		245
<i>C. odorata</i>	506	0	1	0	3	0		72
<i>D. guianense</i>	402	0	5	0	4	0		58
<i>D. odorata</i>	737	0	3	0	6	1		101
<i>E. odorata</i>	790	1	3	2	8	0		107
<i>H. crepitans</i>	376	0	5	4	5	0		46
<i>H. courbaril</i>	443	0	3	0	5	0		42
<i>J. copaia</i>	1047	0	4	3	9	0		132
<i>P. pendula</i>	503	0	1	0	3	0		84
<i>P. riedellii</i>	414	0	1	1	3	0		79
<i>P. minima</i>	423	0	5	0	8	0		51
<i>S. lutea</i>	347	0	3	0	4	0		53
<i>T. serratifolia</i>	499	0	0	1	4	0		62
<i>T. guianensis</i>	343	0	0	0	1	0		56
<i>T. altissima</i>	1736	3	7	1	24	0		195
Erro de Omissão (%)	30	100	100	100	99	100	...	88
Erro de Comissão (%)	92	100	96	90	93	100	...	95

Horta (2006), ao realizar levantamento dos solos e ocupação da superfície do município de Nazareno, Minas Gerais, averiguou que, utilizando imagem com resolução espacial de 30 m, houve falha na classificação de cerca de 75% das voçorocas, alocadas nas classes de vegetação arbustiva, cultura, pastagens ou campos naturais. Mesmo utilizando resolução espacial de 5 m no presente trabalho, a classificação não apresentou acurácia adequada, indicando que seria necessária utilização de imagem de melhor resolução.

A caracterização espectral em nível foliar, como realizado na Mata Atlântica por Miyoshi (2016), por meio de espectroscopia de amostras de vegetação, pode contribuir positivamente para a melhor acurácia na classificação da vegetação nativa. A utilização de Veículos Aéreos Não Tripulados (VANT) também é recomendada para esse tipo de estudo, visto que a resolução espacial pode chegar a 10 cm (CÂNDIDO et al., 2015).

#### 4. Conclusões

Tanto a técnica de máxima verossimilhança quanto a resolução espacial da imagem utilizada não foram apropriadas para a classificação supervisionada de vegetação nativa. 13 das 20 espécies selecionadas no estudo de campo não foram identificadas pelo método, sendo que, entre as 7 identificadas, 2 foram subestimadas e as restantes superestimadas.

Estudos utilizando caracterização espectral em nível foliar e imagens com maior resolução são recomendados para aumentar a acurácia da classificação.

## Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e à Universidade Federal de Lavras (UFLA) pelo apoio a esta pesquisa.

## Referências Bibliográficas

Beltrame, A. M. K.; Jardini, M. G. M.; Jacobsen, R. M.; Quintanilha, J. A. Análise da vegetação em faixas de servidão de linhas de transmissão de energia elétrica. In: Simpósio Brasileiro de Sensoriamento Remoto (SBSR), 13, 2007, Florianópolis. **Anais...** São José dos Campos: INPE, 2007. Disponível em: <martes.sid.inpe.br/col/dpi.inpe.br/sbsr@80/2006/11.15.11.46/doc/5579-5586.pdf>. Acesso em: 31 out. 2016.

Cândido, A.K.A.A; Silva, N.M.; Paranhos Filho, A.C. Imagens de Alta Resolução Espacial de Veículos Aéreos Não Tripulados (VANT) no Planejamento do Uso e Ocupação do Solo. **Anuário do Instituto de Geociências – UFRJ**, v. 38, n. 1, p. 147-156, 2015.

Celinski, T.M.; Zimback, C.R.L. Discriminação de classes de cobertura vegetal utilizando técnicas de classificação digital de imagens de sensoriamento remoto. **Energia na Agricultura**, v. 25, n. 1, p. 152-170, 2010.

Chagas, C.S.; Vieira, C.A.O.; Fernandes Filho, E.I. Comparison between artificial neural networks and maximum likelihood classification in digital soil mapping. **Revista Brasileira de Ciência do Solo**, v. 37, n. 2, p. 339-351, 2013.

Horta, I.M.F. **Levantamento dos solos e ocupação da superfície do município de Nazareno, MG**. 2006. 74 p. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Lavras, Lavras. 2006.

Jensen, J.R. **Introductory Digital Image Processing: A Remote Sensing Perspective**. 3rd Edition, Upper Saddle River: Prentice-Hall, 2005. 526 p.

Martins F.R. **Estrutura de uma floresta mesófila**. Campinas: Ed. UNICAMP, 1991. 246 p.

Meirelles, M.S.P.; Camara Netto, G. Almeida, C.M. **Geomática - Modelos e Aplicações Ambientais**. 1. ed. Brasília: Embrapa Informação Tecnológica, 2007. 593 p.

Miyoshi, G.T. **Caracterização espectral de espécies de Mata Atlântica de Interior em nível foliar e de copa**. 2015. 142 p. Dissertação (Mestrado em Ciências Cartográficas) - Universidade Estadual Paulista, Presidente Prudente. 2016.

Moras Filho, L.O. **Método de inferência espacial para o ordenamento territorial de municípios de pequeno porte: o caso de Ijaci, MG**. 2015. 95 p. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Lavras, Lavras. 2015.

Moura, A.C.M. **Geoprocessamento na gestão e planejamento urbano**. Belo Horizonte. Ed. da Autora, 2003. 294 p.

Novo, E.M.L de M. **Sensoriamento remoto: princípios e aplicações**. 2.ed. São Paulo: Edgard Blucher, 1998. 308p.

Novo, E.M.L.M. **Sensoriamento remoto: princípios e aplicações**. São Paulo: Ed. Edgard Blucher Ltda., 1992. 307 p.

Ortiz, J.L.; Freitas, M.I.C. Mapeamento do uso da terra, vegetação e impactos ambientais por meio de sensoriamento remoto e geoprocessamento. **Geociências**, v. 24, n. 1, p. 91-104, 2005.

Ribeiro, R.M.P. **Avaliação de métodos de classificação de imagens IKONOS II para o mapeamento da cobertura terrestre**. 2004. 53 p. Dissertação (Mestrado em Ciência Florestal) - Universidade Federal de Viçosa, Viçosa. 2004.